

# Phân Tích Dữ Liệu Thực Tế với Python

## Bài 10.2: Xây Dựng Dashboard với Plotly



Quang-Khai Tran, Ph.D  
CyberLab, 04/2023



(Ảnh: Internet)



# Nội dung



1. **Giới thiệu thư viện Plotly**
2. **Sử dụng Plotly**
  - Các chế độ Online vs. Offline
  - Plotly Express
  - Plotly Graph\_Object
  - Hiển thị bản đồ
3. **Demo: Tạo dashboard cho phân tích EDA**
4. **Bài tập & Thảo Luận**

# Phần 1. Giới thiệu thư viện Plotly



- ❖ Do hãng Dash Enterprise xây dựng và chia sẻ (open source)



<https://plotly.com/>



- ❖ Plotly là thư viện cho phép tạo ra các biểu đồ có tính tương tác mạnh mẽ:
  - Plotly hiển thị hình ảnh dựa trên nền tảng các browsers
  - Có thể kết hợp tốt với các thư viện khác trong Python (Numpy, Pandas)
  - Các use-cases: statistical, financial, geographic, scientific, và 3-dimensional



## Plotly: The front end for ML and data science models

<https://plotly.com>

Dash apps go where Tableau and PowerBI cannot: NLP, object detection, predictive analytics, and more

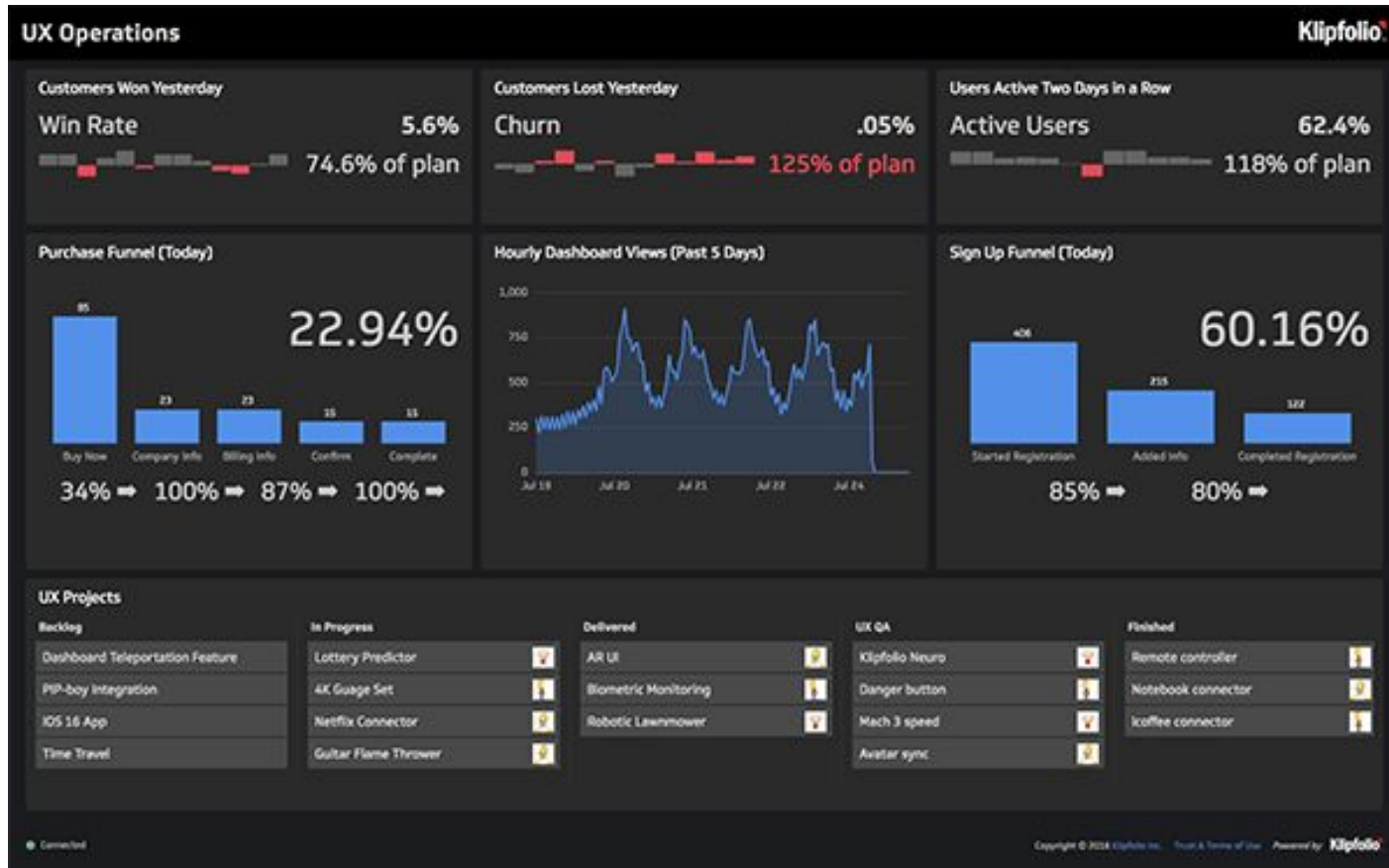
- ❖ Plotly thường được kết hợp với thư viện Dash để tạo dashboard cho webapp:
  - Thường sử dụng nhiều trong các dự án DA có yêu cầu cao về trình bày
  - Giúp tạo ra các analytic web apps (còn gọi là Dash apps)
  - Đặc biệt hướng đến trực quan hóa phân tích dữ liệu và machine learning
  - Tuy nhiên, để vận dụng tốt Dash cần nắm vững HTML và CSS (vì được xây trên nền Plotly.js and React.js)

# Một số ví dụ

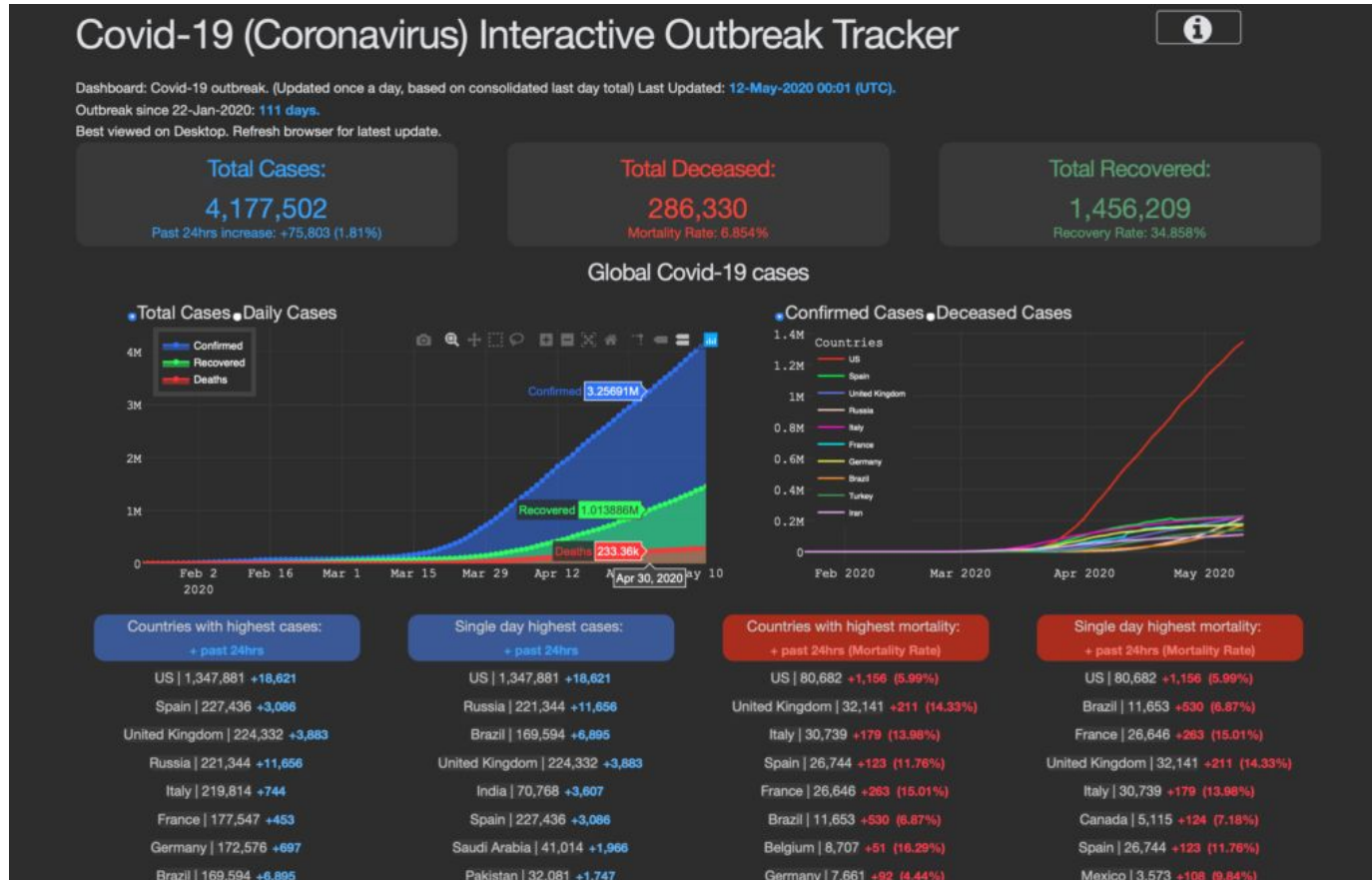




# Một số ví dụ

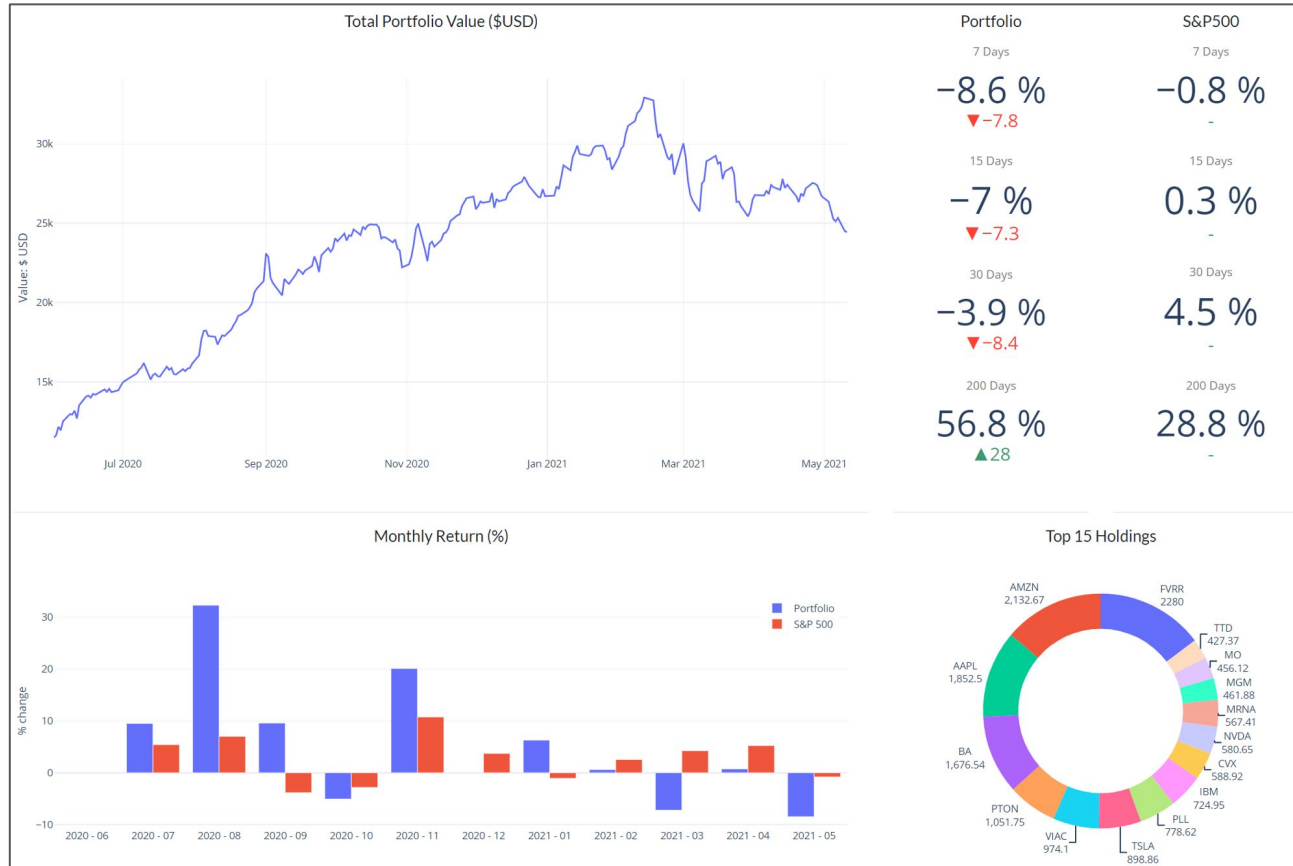


# Một số ví dụ

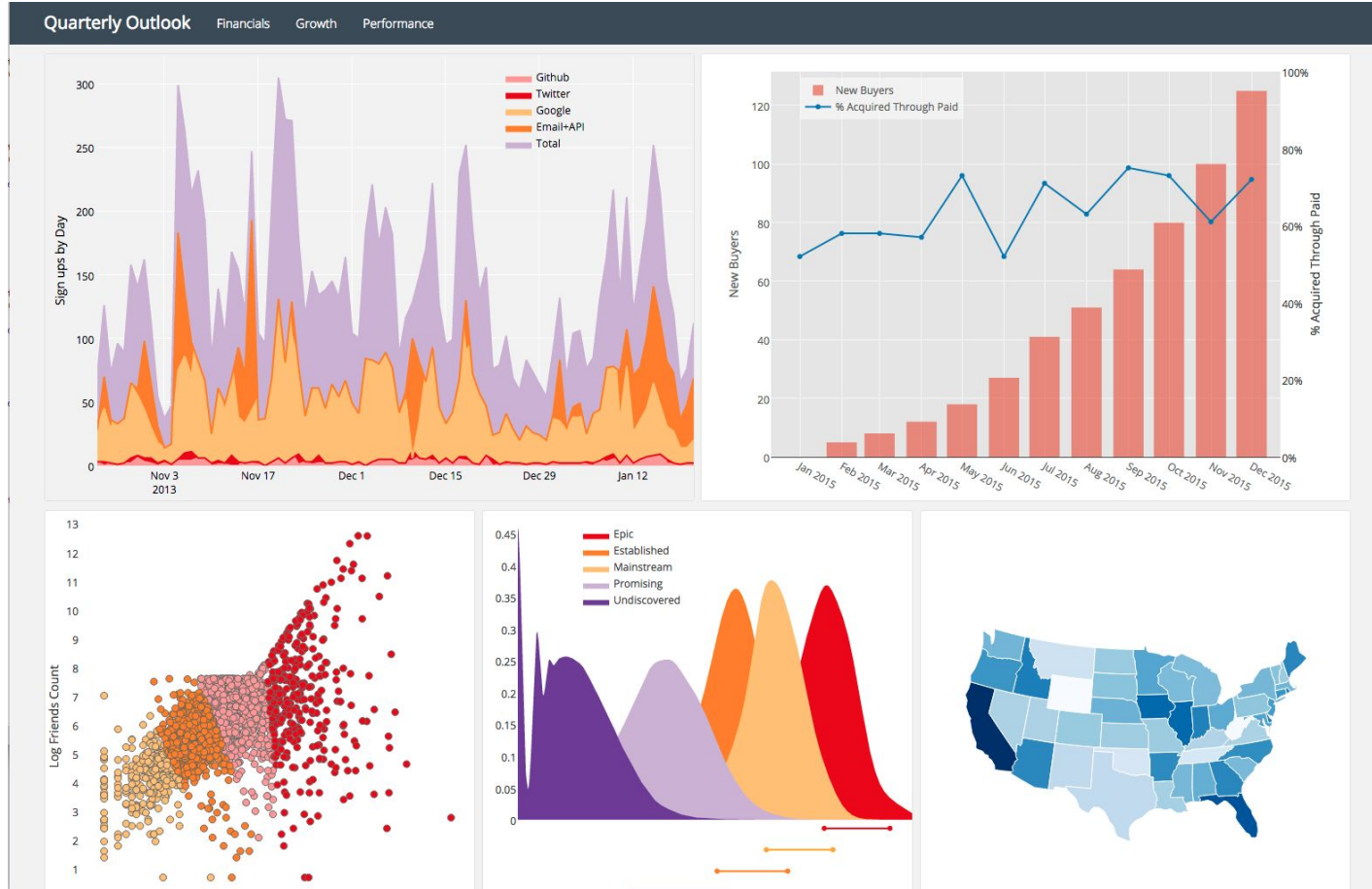




# Một số ví dụ



# Một số ví dụ



## ❖ Cài đặt plotly và jupyter-dash

```
In [3]: %pip install plotly
```

```
Collecting plotly
  Downloading plotly-5.3.1-py2.py3-none-any.whl (23.9 MB)
    |████████████████████| 23.9 MB 28 kB/s eta 0:00:01
Collecting tenacity>=6.2.0
  Downloading tenacity-8.0.1-py3-none-any.whl (24 kB)
Requirement already satisfied: six in /home/tqkhai/Working/Tools/anaconda3/lib/python3.8/site-packages (from plotly) (1.15.0)
Installing collected packages: tenacity, plotly
Successfully installed plotly-5.3.1 tenacity-8.0.1
Note: you may need to restart the kernel to use updated packages.
```

```
In [4]: %pip install jupyter-dash
```

```
Collecting jupyter-dash
  Downloading jupyter_dash-0.4.0-py3-none-any.whl (20 kB)
Requirement already satisfied: flask in /home/tqkhai/Working/Tools/anaconda3/lib/python3.8/site-packages (from jupyter-dash) (1.1.2)
Collecting ansi2html
  Downloading ansi2html-1.6.0-py3-none-any.whl (14 kB)
Collecting dash
  Downloading dash-2.0.0-py3-none-any.whl (7.3 MB)
    |████████████████████| 7.3 MB 535 kB/s eta 0:00:01
Requirement already satisfied: ipython in /home/tqkhai/Working/Tools/anaconda3/lib/python3.8/site-packages (from jupyter-dash) (7.26.0)
Collecting retrying
  Downloading retrying-1.3.3.tar.gz (10 kB)
```

## Một số nguồn tham khảo chính:

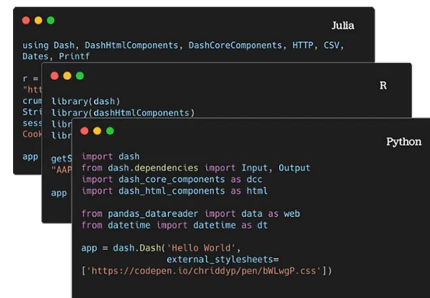
- ❖ <https://plotly.com/python/>
- ❖ <https://plotly.com/python/basic-charts/>
- ❖ <https://www.justintodata.com/python-interactive-dashboard-with-plotly-dash-tutorial/>
- ❖ <https://www.geeksforgeeks.org/python-plotly-tutorial/>

## Một số tài liệu Tiếng Việt:

- ❖ <https://codelearn.io/sharing/thu-vien-plotly-trong-python-la-gi>
- ❖ <https://ichi.pro/vi/vuot-len-tren-ngan-nap-plotly-express-hien-chap-nhan-du-lieu-dang-rong-va-dang-hon-hop-208648979937425>

## Phần 2. Sử dụng Plotly

- 2.1. Các chế độ Online vs. Offline
- 2.2. Các modules chính
  - 2.1. Express và Graph\_Objects
  - 2.2. Dash
- 2.3. Hiển thị bản đồ



```
Julia
using Dash, DashHtmlComponents, DashCoreComponents, HTTP, CSV,
Dates, Printf

r =
R
library(dash)
library(dashHtmlComponents)

Python
import dash
from dash.dependencies import Input, Output
import dash_core_components as dcc
import dash_html_components as html

from pandas_datareader import data as web
from datetime import datetime as dt

app = dash.Dash('Hello World',
external_stylesheets=
['https://codepen.io/chriddyp/pen/BNLwgP.css'])
```

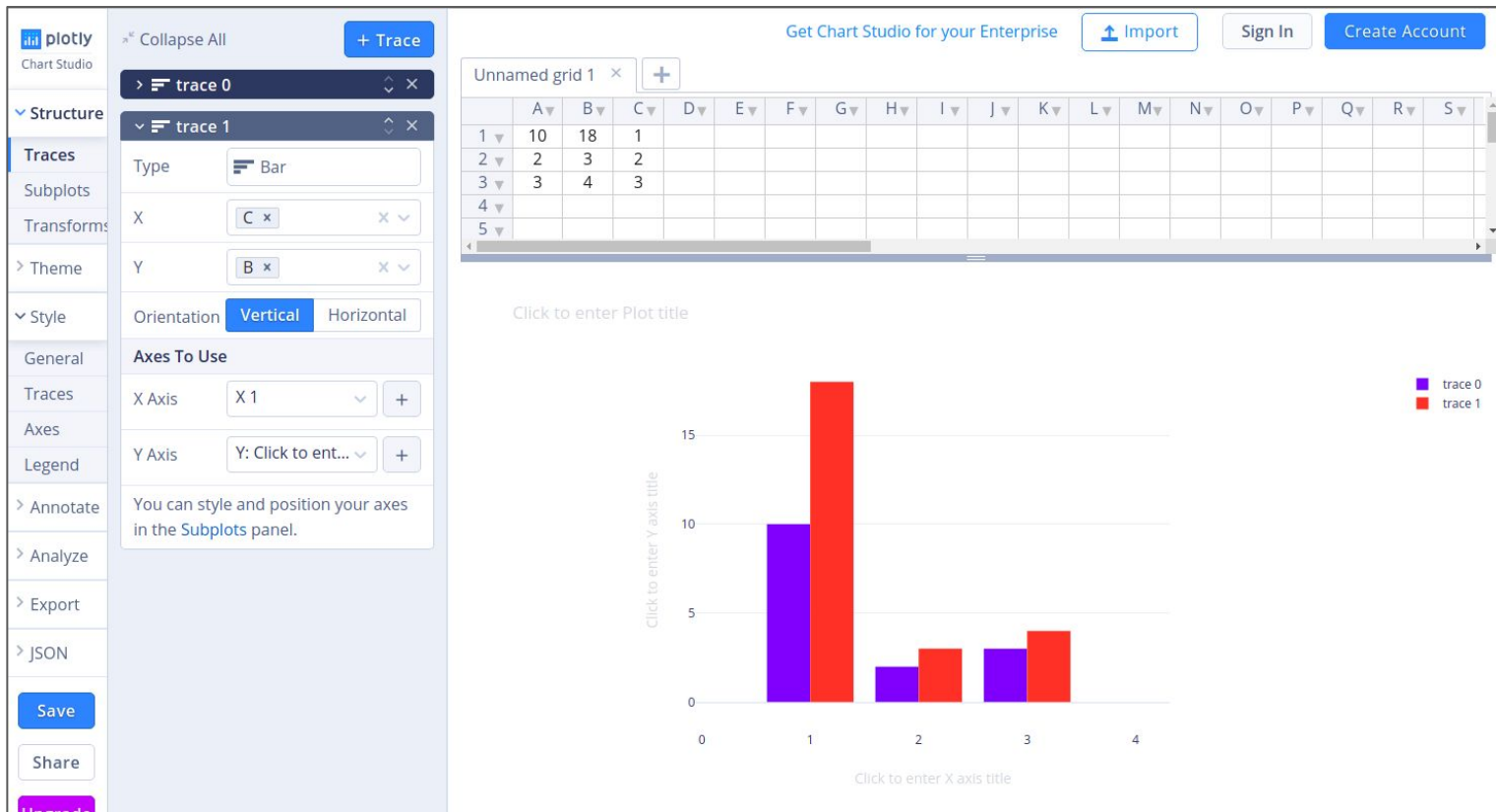
## 2.1 Các chế độ Online vs Offline của Plotly

- ❖ Online:
  - Các hình vẽ và dữ liệu liên quan được lưu trên server của Plotly
  - Link: <https://chart-studio.plotly.com/create/#/>
  - Hướng dẫn: <https://plotly.com/chart-studio-help/excel/dashboard/>
  
- ❖ Offline:
  - Cho phép tạo ra các biểu đồ offline và lưu trên máy tính
  - Lưu thành file html: `plotly.offline.plot()`
  - Hiển thị trên Jupyter Notebook: `plotly.offline.iplot()`



## 2.1 Các chế độ Online vs Offline của Plotly

Online: <https://chart-studio.plotly.com/create/#/>





plotly | Express



plotly.graph\_objects

## 2.2 Các modules chính: Express

### ❖ `plotly.express`:

- Là module bậc cao cung cấp các hàm vẽ biểu đồ cơ bản và đơn giản nhất
- Được đề xuất để bắt đầu tìm hiểu về Plotly
- Có thể coi là tương tự như các hàm vẽ với pyplot (plt) trong matplotlib, hoặc seaborn với matplotlib
- Tham khảo: <https://plotly.com/python/plotly-express/>

*plotly.express is to plotly what seaborn is to matplotlib*

## 2.2 Các modules chính: Graph\_Objects

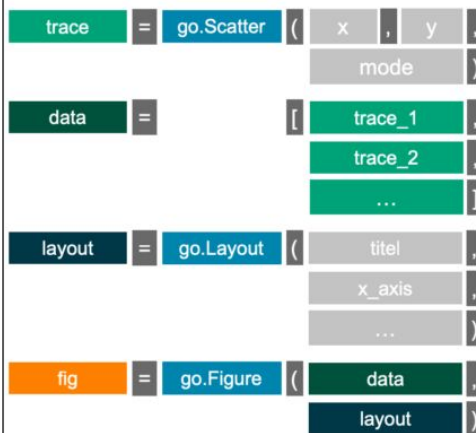
- ❖ **plotly.graph\_objects:**
  - Là module bậc dưới của plotly.express
  - Gồm các hàm tương tác với Plotly.js và React.js
  - Hai classes chính là Figure và FigureWidget (để tương thích với ipywidgets)
  - Có thể coi là tương tự như các axes trong matplotlib
  - Tham khảo: <https://plotly.com/python/graph-objects/>
  
- ❖ Full API Reference:  
<https://plotly.com/python-api-reference/>

## 2.2 Các modules chính: Graph\_Objects

### Cấu trúc của graph\_objects

- ❖ **trace**: thông số trên biểu đồ
- ❖ **data**: các trace(s)
- ❖ **layout**: các yếu tố trình bày
- ❖ **figure**: đối tượng biểu đồ

### STRUCTURE OF A PLOTLY CHART



```
trace = go.Scatter(
    x = df.arr_delay,
    y = df.dep_delay,
    mode = 'markers')

data = [trace]

layout=go.Layout(title="Departure Delay vs
    Arrival Delay",
    xaxis={'title':'Arrival Delay'},
    yaxis={'title':'Departure Delay'})

fig = go.Figure(data=data, layout=layout)
iplot(fig)
```



STATWORX

## 2.2 Các modules chính: dash

### ❖ Dash là một thư viện riêng để phát triển các applications trên web

- Để sử dụng dash cần hiểu về html
- Tham khảo: <https://dash.plotly.com/>



```
import dash
import dash_core_components as dcc
import dash_html_components as html

app = dash.Dash()
app.layout = html.Div([
    dcc.Graph(figure=fig)
])

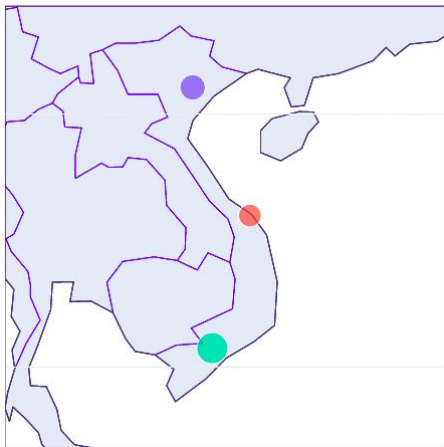
app.run_server(debug=True, use_reloader=False)
```



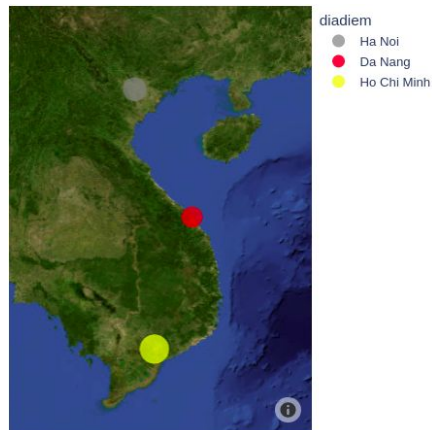
## 2.3 **Hiện thị bản đồ**

Tham khảo: <https://plotly.com/python/maps/>

`px.scatter_geo()`



`px.scatter_mapbox()`



`go.Scattergeo()`

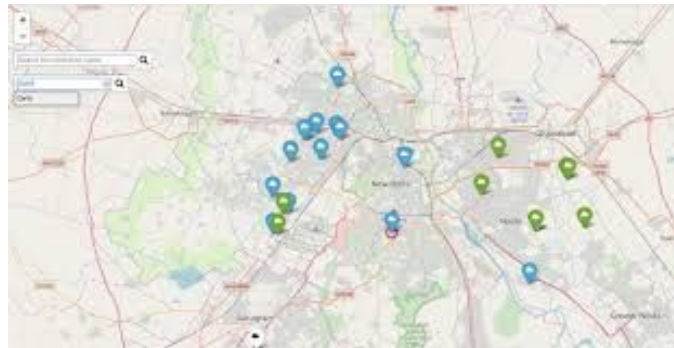


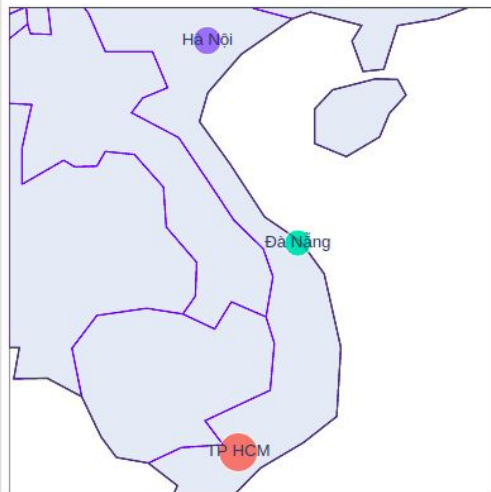
## 2.4 **Hiển thị bản đồ - Folium**

### Sử dụng bản đồ với thư viện Folium

Tham khảo 1: <https://phamdinhhkhanh.github.io/2019/10/22/googleHeatmap.html>

Tham khảo 2: <https://python-visualization.github.io/folium/quickstart.html>





Chọn Thành Phố

Tất Cả

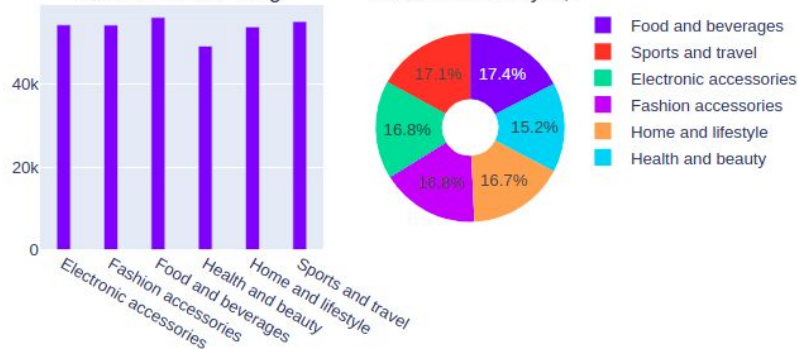
Hà Nội

TP HCM

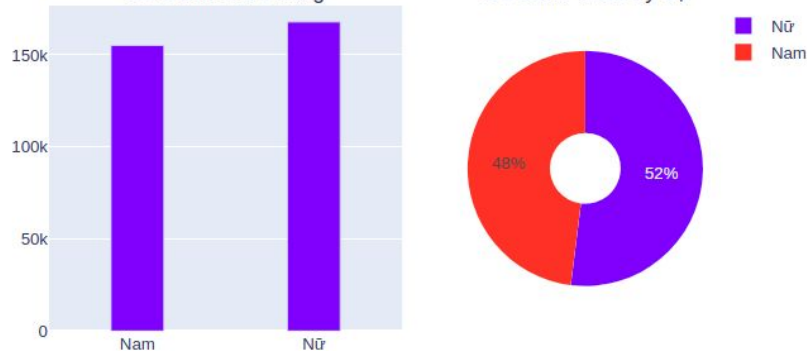
Đà Nẵng

THÀNH PHỐ: TẤT CẢ

BIỂU ĐỒ SO SÁNH DOANH THU THEO: Product line  
So Sánh Theo Tổng So Sánh Theo Tỷ Lệ



BIỂU ĐỒ SO SÁNH DOANH THU THEO: Gender  
So Sánh Theo Tổng So Sánh Theo Tỷ Lệ



1. Chuẩn bị cho đồ án cuối khóa: Xây dựng dashboard để thực hiện phân tích EDA
2. Nâng cao (optional):
  - Thực hiện phân tích EDA về dữ liệu giá nhà hoặc NYCFlights bằng Plotly
  - Tạo dashboards